REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
04-04-2011	Final Technical	01-Mar-2008 to 31-Dec-2010

4. TITLE AND SUBTITLE
N00014 08 1 0882 Algebraic-Topological Structures for Hidden Modes

5a. CONTRACT NUMBER

5b. GRANT NUMBER
N00014-08-1-0882

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Dr. Yuliy Baryshnikov

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Lucent Technologies (Alcatel-Lucent / Bell Laboratories)
600 Mountain Ave
Murray Hill, NJ 07974

8. PERFORMING ORGANIZATION REPORT NUMBER
N00014-08-1-0882 FINAL

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
875 North Randolph Street
Arlington, VA 22203-1995

10. SPONSOR/MONITOR'S ACRONYM(S)
ONR

11. SPONSORING/MONITORING AGENCY REPORT NUMBER

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for public release, distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT
In this report we introduce and present initial findings in the novel, topology based, approach to data analysis. The tools we develop are generally covariant and therefore independent of the particular coordinate system, or its semantics.

15. SUBJECT TERMS
Algebraic Topological, Data Analysis

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
U	U	U	SAR	23	Brian Moore
					19b. TELEPONE NUMBER *(Include area code)* (908) 582 - 3970

Technical Report

N00014 08 1 0882 Algebraic-Topological Structures for Hidden Modes

Dr. Yuliy Baryshnikov
ymb@research.bell-labs.com
Mathematical and Algorithmic Sciences
Bell Laboratories, Alcatel/Lucent,
Murray Hill, NJ; phone/fax 908-582-3839

Administrative/business contact: Brian Moore, Bell Labs Grant Manager, moorebf@alcatel-lucent.com, 1D-470, 600-700 Mountain Avenue PO Box 636, Murray Hill, NJ 07974-0636, phone 908-582-3970, fax 908-582-3104.

Duration of effort: March 1, 2008 - December 31, 2010

Abstract: In this report we introduce and present initial findings in the novel, topology based, approach to data analysis. The tools we develop are generally covariant and therefore independent of the particular coordinate system, or its semantics.

This report outlines results obtained in [2] and [3].

CONTENTS

1. Euler Integration in o-minimal setting

1.1. Introduction. Integration with respect to Euler characteristic is a homomorphism $\int_X \cdot d\chi : CF(X) \to \mathbb{Z}$ from the ring of constructible functions $CF(X)$ ("tame" integer-valued functions on a topological space X) to the integers $\mathbb{Z}$. It is a topological integration theory which uses as a measure the venerable Euler characteristic χ. Euler characteristic, suitably defined, satisfies the fundamental property of a measure:

$$\chi(A \cup B) = \chi(A) + \chi(B) - \chi(A \cap B), \tag{1.1}$$

for A and B "tame" subsets of X. We extend the theory to $\mathbb{R}$-valued integrands and demonstrate its utility in managing incomplete data in, e.g., sensor networks.

1.2. Constructible integrands. Because the Euler characteristic is only finitely additive, one must continually invoke the word "tame" to ensure that χ is well-defined. One means by which to do so it via an O-MINIMAL STRUCTURE **[39]**, a sequence $\mathcal{O} = (\mathcal{O}_n)$ of Boolean algebras of subsets of $\mathbb{R}^n$ satisfying a small list of axioms: closure under products, closure under projections, and finite decompositions in $\mathcal{O}_1$. Elements of $\mathcal{O}$ are called DEFINABLE sets and these are "tame" for purposes of integration theory. Examples of o-minimal structures include (1) piecewise-linear sets;[1] (2) semi-algebraic sets; and (3) globally subanalytic sets.

Definable functions between spaces are those whose graphs are in $\mathcal{O}$. For X and Y definable spaces, let $\mathsf{Def}(X, Y)$ denote the class of compactly supported definable functions $h : X \to Y$, and fix as a convention $\mathsf{Def}(X) = \mathsf{Def}(X, \mathbb{R})$. Let $CF(X) = \mathsf{Def}(X, \mathbb{Z}) \subset \mathsf{Def}(X, \mathbb{R})$ denote the ring of CONSTRUCTIBLE FUNCTIONS: compactly supported $\mathbb{Z}$-valued functions all of whose level sets are definable. Note that in general, definable functions (even definable 'homeomorphisms') are not necessarily continuous.

We briefly recall the theory of Euler integration, established as an integration theory in the constructible setting in **[28, 36, 35, 38]** and anticipated by a combinatorial version in **[6, 21, 22, 34]**. Fix an o-minimal structure $\mathcal{O}$ on a space X. The geometric Euler characteristic is the function $\chi : \mathcal{O} \to \mathbb{Z}$ which takes a definable set $A \in \mathcal{O}$ to $\chi(A) = \sum_i (-1)^i \dim H_i^{BM}(A; \mathbb{R})$, where H_*^{BM} is the Borel-Moore homology (equivalently, singular compactly supported cohomology)

---

[1] Some authors require an o-minimal structure to contain algebraic curves, eliminating this particular example.

of A. This also has a combinatorial definition: if A is definably homeomorphic to a finite disjoint union of (open) simplices $\coprod_j \sigma_j$, then $\chi(A) = \sum_j (-1)^{\dim \sigma_j}$. Algebraic topology asserts that χ is independent of the decomposition into simplices. The Mayer-Vietoris principle asserts that χ is a measure (or 'valuation') on $\mathcal{O}$, as expressed in Eqn. [1.1].

The EULER INTEGRAL is the pushforward of the trivial map $X \mapsto \{pt\}$ to $\int_X d\chi : CF(X) \to CF(\{pt\}) \cong \mathbb{Z}$ satisfying $\int_X \mathbb{1}_A \, d\chi = \chi(A)$ for $\mathbb{1}_A$ the characteristic function over a definable set A. From the definitions and a telescoping sum one easily obtains:

$$\int_X h \, d\chi = \sum_{s=-\infty}^{\infty} s\chi\{h = s\} = \sum_{s=0}^{\infty} \chi\{h > s\} - \chi\{h < -s\}. \tag{1.2}$$

Because the Euler integral is a pushforward, any definable map $F : X \to Y$ induces $F_* : CF(X) \to CF(Y)$ satisfying $\int_X h \, d\chi = \int_Y F_* h \, d\chi$. Explicitly,

$$F_* h(y) = \int_{F^{-1}(y)} h \, d\chi, \tag{1.3}$$

as a manifestation of the Fubini Theorem.

The Euler integral has been found to be an elegant and useful tool for explaining properties of algebraic curves [8] and stratified Morse theory [37, 9], for reconstructing objects in integral geometry [35], for target counting in sensor networks [1], and as an intuitive basis for the more general theory of motivic integration [12, 13].

1.3. Real-valued integrands. We extend the definition of Euler integration to $\mathbb{R}$-valued integrands in $\mathrm{Def}(X)$ via step-function approximations.

1.3.1. *A Riemann-sum definition.*

DEFINITION 1. Given $h \in \mathrm{Def}(X)$, define:

$$\int_X h \, \lfloor d\chi \rfloor = \lim_{n \to \infty} \frac{1}{n} \int_X \lfloor nh \rfloor d\chi. \tag{1.4}$$

$$\int_X h \, \lceil d\chi \rceil = \lim_{n \to \infty} \frac{1}{n} \int_X \lceil nh \rceil d\chi. \tag{1.5}$$

We establish that these limits exist and are well-defined, though not equal.

LEMMA 2. *Given an affine function $h \in \mathsf{Def}(\sigma)$ on an open k-simplex σ,*

$$\int_\sigma h\,\lfloor d\chi \rfloor = (-1)^k \inf_\sigma h\,; \int_\sigma h\,\lceil d\chi \rceil = (-1)^k \sup_\sigma h. \tag{1.6}$$

This integration theory is robust to changes in coordinates.

LEMMA 3. *Integration on $\mathsf{Def}(X)$ with respect to $\lfloor d\chi \rfloor$ and $\lceil d\chi \rceil$ is invariant under the right action of definable bijections of X.*

LEMMA 4. *The limits in Definition 1 are well-defined.*

Integrals with respect to $\lfloor d\chi \rfloor$ and $\lceil d\chi \rceil$ are related to total variation (in the case of compactly supported continuous functions).

COROLLARY 5. If M is a 1-dimensional manifold and $h \in \mathsf{Def}(M)$ is continuous, then

$$\int_M h\,\lfloor d\chi \rfloor = -\int_M h\,\lceil d\chi \rceil = \frac{1}{2}\mathsf{totvar}(h). \tag{1.7}$$

This result generalizes greatly via Morse theory: see Corollary 11. One notes that $\lfloor d\chi \rfloor$ and $\lceil d\chi \rceil$ give integrals which are conjugate in the following sense.

LEMMA 6.

$$\int h\lceil d\chi \rceil = -\int -h\lfloor d\chi \rfloor. \tag{1.8}$$

The temptation to cancel the negatives must be resisted: see Lemma 12 below.

1.3.2. *Computation.* Definition 1 has a Riemann-sum flavor which extends to certain computational formulae. The following is a definable analogue of Eqn. [1.2].

PROPOSITION 7. *For $h \in \mathsf{Def}(X)$,*

$$\int_X h\,\lfloor d\chi \rfloor = \int_{s=0}^{\infty} \chi\{h \geq s\} - \chi\{h < -s\}\,ds \tag{1.9}$$

$$\int_X h\,\lceil d\chi \rceil = \int_{s=0}^{\infty} \chi\{h > s\} - \chi\{h \leq -s\}\,ds. \tag{1.10}$$

It is not true that $\int_X h\,\lfloor d\chi \rfloor = \int_0^\infty s\chi\{h = s\}ds$: the proper Lebesgue generalization of Eqn. [1.2] is the following:

PROPOSITION 8. *For* $h \in \mathsf{Def}(X)$,

$$\int_X h \lfloor d\chi \rfloor = \lim_{\epsilon\to 0^+} \frac{1}{\epsilon}\int_{\mathbb{R}} s\,\chi\{s \le h < s+\epsilon\}\,ds \tag{1.11}$$

$$\int_X h \lceil d\chi \rceil = \lim_{\epsilon\to 0^+} \frac{1}{\epsilon}\int_{\mathbb{R}} s\,\chi\{s < h \le s+\epsilon\}\,ds. \tag{1.12}$$

1.3.3. *Morse theory.* One important indication that the definition of $\int \lfloor d\chi \rfloor$ is correct for our purposes is the natural relation to Morse theory: the integrals with respect to $\lfloor d\chi \rfloor$ and $\lceil d\chi \rceil$ are Morse index weighted sums of critical values of the integrand. This is a localization result, reducing from an integral over all of X to an integral over an often discrete set of critical points.

Recall that a C^2 function $h : M \to \mathbb{R}$ on a smooth manifold M is MORSE if all critical points of h are nondegenerate, in the sense of having a nondegenerate Hessian matrix of second partial derivatives. Denote by $\mathcal{C}(h)$ the set of critical points of h. For each $p \in \mathcal{C}(h)$, the MORSE INDEX of p, $\mu(p)$, is defined as the number of negative eigenvalues of the Hessian at p, or, equivalently, the dimension of the unstable manifold $W^u(p)$ of the vector field $-\nabla h$ at p.

Stratified Morse theory **[20]** is a powerful generalization to triangulable spaces, including definable sets with respect to an o-minimal structure **[9, 37]**. We may interpret $\lfloor d\chi \rfloor$ and $\lceil d\chi \rceil$ in terms of the weighted stratified Morse index of the graph of the integrand.

DEFINITION 9. For $X \subset \mathbb{R}^n$ definable and $h \in \mathsf{Def}(X)$, define the co-index of h, $\mathcal{I}^*h$ to be the stratified Morse index of the graph of h, $\Gamma_h \subset X \times \mathbb{R}$, with respect to the projection $\pi : X \times \mathbb{R} \to \mathbb{R}$:

$$(\mathcal{I}^*h)(x) = \lim_{\epsilon' \ll \epsilon \to 0^+} \chi\left(\overline{B_\epsilon(x)} \cap \{h < h(x) + \epsilon'\}\right), \tag{1.13}$$

where $\overline{B_\epsilon(x)}$ is the closed ball of radius ϵ about $x \in X$. The index $\mathcal{I}_*$ is the stratified Morse index with respect to height function $-\pi$: i.e., $\mathcal{I}_*h = \mathcal{I}^*(-h)$ or

$$(\mathcal{I}_*h)(x) = \lim_{\epsilon' \ll \epsilon \to 0^+} \chi\left(\overline{B_\epsilon(x)} \cap \{h > h(x) - \epsilon'\}\right). \tag{1.14}$$

Note that $\mathcal{I}_*, \mathcal{I}^* : \mathsf{Def}(X) \to CF(\overline{X})$, and the restriction of these operators to $CF(X)$ is the identity (every point of a constructible function is a critical point). The two types of integration on $\mathsf{Def}(X)$ correspond to the Morse indices of the graph with respect to the two orientations of the graph axis — the projections π and $-\pi$.

THEOREM 10. *For any continuous* $h \in \mathsf{Def}(X)$,

$$\int_X h\,\lfloor d\chi \rfloor = \int_{\overline{X}} h\mathcal{I}^* h\, d\chi \quad ; \quad \int_X h\,\lceil d\chi \rceil = \int_{\overline{X}} h\mathcal{I}_* h\, d\chi. \tag{1.15}$$

COROLLARY 11. If h is a Morse function on a closed n-manifold M, then:

$$\int_M h\,\lfloor d\chi \rfloor = \sum_{p\in C(h)} (-1)^{n-\mu(p)} h(p); \tag{1.16}$$

$$\int_M h\,\lceil d\chi \rceil = \sum_{p\in C(h)} (-1)^{\mu(p)} h(p). \tag{1.17}$$

From this, one sees clearly that the relationship between $\lfloor d\chi \rfloor$ and $\lceil d\chi \rceil$ is regulated by Poincaré duality. For example, on continuous definable integrands over an n-dimensional manifold M,

$$\int_M h\,\lceil d\chi \rceil = (-1)^n \int_M h\,\lfloor d\chi \rfloor. \tag{1.18}$$

The generalization from continuous to general definable integrands is simple, but requires weighting $\mathcal{I}^* h$ by h directly. To compute $\int_X h\lfloor d\chi \rfloor$, one integrates the weighted co-index

$$\lim_{\epsilon' \ll \epsilon \to 0^+} h(x+\epsilon')\chi\left(\overline{B_\epsilon(x)} \cap \{h < h(x)+\epsilon'\}\right) \tag{1.19}$$

with respect to $d\chi$.

Corollary 11 can also be proved directly using classical handle-addition techniques or in terms of the Morse complex, using the fact that the restriction of the integrand to each unstable manifold of each critical point is unimodal with a unique maximum at the critical point. It is also possible to express the stratified Morse index — and thus the integral here considered — in terms of integration against a characteristic cycle, *cf.* **[20, 37]**.

One final means of illustrating Corollary 11 is to use a deformation argument. Let h be smooth on X and ϕ_t be the flow of $-\nabla h$. Then the integral is invariant under the action of ϕ_t on h; yet the limiting function $h_\infty = \lim_{t\to\infty} h \circ \phi_t$ is constant on stable manifolds of $-\nabla h$ with values equal to the critical values of h. We have not shown that the limiting function is constructible (this depends on the existence of definable invariant manifolds — we are unaware of relevant results in the literature) and thus do not rely on this method for proof but rather illumination.

1.4. The integral operator. We consider properties of the integral operator(s) on Def(X).

1.4.1. *Linearity.* One is tempted to apply all the standard constructions of sheaf theory (as in **[36, 35]**) to $\int_X : \mathrm{Def}(X) \to \mathbb{R}$. However, our formulation of the integral on Def(X) has a glaring disadvantage.

LEMMA 12. *$\int_X : \mathrm{Def}(X) \to \mathbb{R}$ (via $\lfloor d\chi \rfloor$ or $\lceil d\chi \rceil$) is not a homomorphism for* $\dim X > 0$.

This loss of functoriality can be seen as due to the fact that $\lfloor f + g \rfloor$ agrees with $\lfloor f \rfloor + \lfloor g \rfloor$ only up to a set of Lebesgue measure zero, not χ-measure zero. The nonlinear nature of the integral is also clear from Eqn. [1.15], as Morse data is non-additive.

1.4.2. *The Fubini Theorem.* In one sense, the change of variables formula trivializes (Lemma 3). The more general change of variables formula encapsulated in the Fubini theorem does not, however, hold for non-constructible integrands.

COROLLARY 13. The Fubini theorem fails on Def(X) in general.

Fubini holds when the map respects fibers.

THEOREM 14. *For $h \in \mathrm{Def}(X)$, let $F : X \to Y$ be definable and h-preserving (h is constant on fibers of F). Then $\int_Y F_*h\lfloor d\chi \rfloor = \int_X h\lfloor d\chi \rfloor$, and $\int_Y F_*h\lceil d\chi \rceil = \int_X h\lceil d\chi \rceil$.*

COROLLARY 15. For $h \in \mathrm{Def}(X)$, $\int_X h = \int_{\mathbb{R}} h_*h$. In other words,

$$\int_X h \lfloor d\chi \rfloor = \int_{\mathbb{R}} s\, \chi\{h = s\} \lfloor d\chi \rfloor, \tag{1.20}$$

and likewise for $\lceil d\chi \rceil$.

1.4.3. *Continuity.* Though the integral operator is not linear on Def(X), it does retain some nice properties. All properties below stated for $\int \lfloor d\chi \rfloor$ hold for $\int \lceil d\chi \rceil$ via duality.

LEMMA 16. *The integral $\int \lfloor d\chi \rfloor : \mathrm{Def}(X) \to \mathbb{R}$ is positively homogeneous.*

Integration is not continuous on Def(X) with respect to the C^0 topology. An arbitrarily large change in $\int h\lfloor d\chi \rfloor$ may be effected by small changes to h on a (large) finite point set. In some situations the "complexity" of the definable functions can be controlled in a way sufficient to ensure continuity.

One example arises in the semialgebraic category. Fix a (finite) semialgebraic stratification $\mathcal{S}$ of a compact definable X, and consider definable semialgebraic functions *algebraic* with respect to this stratification (that is such that the restriction of the function to any stratum $S \in \mathcal{S}$ is a polynomial P_S). The resulting linear space (filtered by the subspaces of polynomials of bounded degree) can be equipped with the structure of a Banach space, by completing the family of seminorms $||P||_{S,\mathbf{k}} = \max_{S\in\mathcal{S}} ||P_S||_{C^n}$, where $n = \dim X$. Then $\int_X \cdot \lfloor d\chi \rfloor$ becomes a continuous (non-linear) functional on this Banach space. The proof results, essentially, from the Bézout theorem (mimicking Thom-Milnor theory): the total number of critical points graph of a polynomial of degree D on a fixed semi-algebraic set is bounded by $O(D^n)$. The generalization to increasing (refined) stratifications is straightforward.

Integration itself defines a natural topology for $\mathsf{Def}(X)$ on which integration is continuous. Define the L^1 ϵ-neighborhood of $h \in \mathsf{Def}(X)$ as the intersection of the C^0 ϵ-neighborhood (definable functions with ϵ-close graphs) with those functions $g \in \mathsf{Def}(X)$ satisfying $\left|\int_X f - g \lfloor d\chi \rfloor\right| < \epsilon$. This provides a basis for an L^1 topology on $\mathsf{Def}(X)$. As a consequence of Lemma 6, the definition is independent of the use of $\lfloor d\chi \rfloor$ or $\lceil d\chi \rceil$.

The interested reader may speculate on other function space topologies on $\mathsf{Def}(X)$.

1.4.4. *Duality and links.* There is an integral transform on $CF(X)$ that is the analogue of Poincaré-Verdier duality [37]. It extends seamlessly to integrals on $\mathsf{Def}(X)$ by means of the following definition.

DEFINITION 17. The DUALITY OPERATOR is the integral transform $\mathcal{D} : CF(X) \to CF(X)$ given by

$$\mathcal{D}h(x) = \lim_{\epsilon \to 0^+} \int_X h \mathbf{1}_{B_\epsilon(x)} d\chi, \tag{1.21}$$

where B_ϵ is an open metric ball of radius ϵ.

We extend the definition to $\mathcal{D} : \mathsf{Def}(X) \to \mathsf{Def}(X)$ by integrating with respect to $\lfloor d\chi \rfloor$ or $\lceil d\chi \rceil$, interchangeably, via:

LEMMA 18. *$\mathcal{D}h$ is well-defined on* $\mathsf{Def}(X)$ *and independent of whether the integration in (1.21) is with respect to $\lfloor d\chi \rfloor$ or $\lceil d\chi \rceil$.*

For a continuous definable function h on a manifold M, $\mathcal{D}h = (-1)^{\dim M} h$, as one can verify by combining Eqns. [1.9] and [1.21]. This is commensurate with the result of Schapira [36] that $\mathcal{D}$ is an involution on $CF(X)$.

THEOREM 19. *Duality is involutive on* $\mathsf{Def}(X)$: $\mathcal{D} \circ \mathcal{D}h = h$.

One can define related integral transforms. For example, the LINK of $h \in CF(X)$ is defined as

$$\Lambda h(x) = \lim_{\epsilon \to 0^+} \int_X h \mathbf{1}_{\partial B_\epsilon(x)} d\chi. \tag{1.22}$$

The link of a continuous function on an n-manifold M is multiplication by $1 + (-1)^n$, as a simple computation shows. In general, $\Lambda = \mathrm{Id} - \mathcal{D}$, where Id is the identity operator.

1.5. Integral transforms. Integration with respect to Euler characteristic over $CF(X)$ has a well-defined and well-studied class of integral transforms, expressed beautifully in Schapira's work on inversion formulae for the generalized Radon transform in $d\chi$ [35]. Integral transforms with respect to $\lfloor d\chi \rfloor$ and $\lceil d\chi \rceil$ are similarly appealing, with applications to signal processing as a primary motivation. Examples of interesting definable kernels for integral transforms over Euclidean $\mathbb{R}^n$ include $\|x-y\|$, $\langle x, y\rangle$, and $g(x-y)$ for some g. These evoke Bessel (Hankel) transforms, Fourier transforms, and convolution with g respectively. The choice between $\lfloor d\chi \rfloor$ and $\lceil d\chi \rceil$ makes a difference, of course, but can be amalgamated. Example: for fixed kernel K, one can consider the mixed integral transform $h \mapsto \int_X hK\lfloor d\chi \rfloor - \int_X hK\lceil d\chi \rceil$. In the case of $K(x, \xi) = \langle x, \xi \rangle$, this transform takes $\mathbf{1}_A$ for A compact and convex to the 'width' of A projected to the ξ-axis.

1.5.1. *Convolution.* On a vector space V (or Lie group, more generally), a convolution operator with respect to Euler characteristic is straightforward. Given $f, g \in CF(V)$, one defines

$$(f * g)(x) = \int_V f(t) g(x-t)\, d\chi. \tag{1.23}$$

Convolution behaves as expected in $CF(V)$. By reversing the order of integration, one has immediately that $\int_V f*g\, d\chi = \int_V f\, d\chi \int_V g\, d\chi$. There is a close relationship between convolution and the Minkowski sum, as observed in, *e.g.*, [21]: for A and B convex and closed $\mathbf{1}_A * \mathbf{1}_B = \mathbf{1}_{A+B}$, cf. [38, 36]. Convolution is a commutative, associative operator providing $CF(V)$ with the structure of an (interesting [8]) algebra.

Convolution is well-defined on $\mathrm{Def}(V)$ by integrating with respect to $\lfloor d\chi \rfloor$ or $\lceil d\chi \rceil$. However, the product formula for $\int f*g$ fails in general, since one relies on the Fubini theorem to prove it in $CF(V)$.

1.5.2. *Linearity.* The nonlinearily of the integration operator prevents most straightforward applications of inversion formulae à la Schapira. Fix a kernel $K \in \mathrm{Def}(X \times Y)$ and consider the integral transform $\mathcal{T}_K : \mathsf{Def}(X) \to \mathsf{Def}(Y)$ of the form $(\mathcal{T}_K h)(y) = \int_X h(x)K(x,y)\lfloor d\chi \rfloor(x)$. In general, this operator is non-linear, via Lemma 12. However, some vestige of (positive) linearity survives within CF.

LEMMA 20. *The integral transform $\mathcal{T}_K$ is positive-linear over $CF^+(X) =$* $\mathsf{Def}(X, \mathbb{N})$.

This implies in particular that when one convolves a function $h \in CF^+(\mathbb{R}^n)$ with a smoothing kernel (e.g., a Gaussian) as a means of filtering noise or taking an average of neighboring data points, that convolution may be analyzed one step at a time (decomposing h).

Integral transforms are not linear over all of $CF(X)$, since $\int -h\lfloor d\chi \rfloor \neq -\int h\lfloor d\chi \rfloor$. However, integral transforms which combine $\lfloor d\chi \rfloor$ and $\lceil d\chi \rceil$ compensate for this behavior. Define the measure $[d\chi]$ to be the average of $\lfloor d\chi \rfloor$ and $\lceil d\chi \rceil$:

$$\int_X h[d\chi] = \frac{1}{2}\left(\int_X h\lfloor d\chi \rfloor + \int_X h\lceil d\chi \rceil\right). \tag{1.24}$$

THEOREM 21. *Any integral transform of the form*

$$(\mathcal{T}_K h)(y) = \int_X h(x)K(x,y)[d\chi](x) \tag{1.25}$$

is a linear operator $CF(X) \to \mathsf{Def}(Y)$.

As a simple example, consider the transform with kernel $K(x,\xi) = \langle x, \xi \rangle$. The transform of 1_A with respect to $[d\chi]$ for A compact and convex equals a 'centroid' of A along the ξ-axis: the average of the maximal and minimal values of ξ on ∂A. Note how the dependence on critical values of the integrand on ∂A reflects the Morse-theoretic interpretation of the integral in this case.

Integration with respect to $[d\chi]$ seems suitable only for integral transforms over CF. On a continuous integrand, the integral with respect to $[d\chi]$ either returns zero (cf. the integral of Rota [34]) or else the integral with respect to $\lfloor d\chi \rfloor$, depending on the parity of the $\dim X$, via Eqn. [1.18].

1.6. Applications of definable Euler integration. The Euler calculus on CF is quite useful; the extension to Def deepens this utility and opens new potential applications, of which we highlight a few.

1.6.1. *Sensor networks.* An application of Euler integration over $CF(X)$ to sensor networks problems was initiated in [1]. Consider a space X whose points represent target-counting sensors that scan a workspace W. Target detection is encoded in a SENSING RELATION $\mathcal{S} \subset W \times X$ where $(w, x) \in \mathcal{S}$ iff a target at w is detected by a sensor at x. Assume that sensors count the number of sensed targets, but do not locate or identify the targets. The sensor network therefore induces a TARGET COUNTING FUNCTION $h : X \to \mathbb{N}$ of the form $h = \sum_\alpha 1_{U_\alpha}$, where U_α is the TARGET SUPPORT — the set of sensors which detect target α. Euler integration allows for simple enumeration:

THEOREM 22 ([**1**]). *Assume $h \in CF(X)$ and $\chi(U_\alpha) = N \neq 0$ for all α. Then the number of targets in W is precisely $\frac{1}{N}\int_X h\,d\chi$.*

Since the target count is presented as an integral, it is possible to accurately estimate the answer when the integrand h is known not on all of X (a continuum of sensors being an idealization) but rather on a sufficiently dense grid of sample points (physical sensors in a network).

The $\mathbb{R}$-valued theory aids in establishing expected values of target counts in the presence of confidence measures on sensor readings. Let $\mathcal{N} = \{x_i\}$ denote a discrete set of sensor nodes in $\mathbb{R}^n$, and assume each sensor returns a target count $h(x_i) \in \mathbb{N}$ and a fluctuation measure $c(x_i) \in [0, 1]$ obtained, say, by stability of the reading over a time average. View h as a sampling over $\mathcal{N}$ of the true target count $f = \sum_\alpha 1_{U_\alpha}$. Assume that nodes with fluctuation reading 0 have perfect information ($h = f$ at x_i) and that c correlates with error $|f - h|$. Assume that sensor nodes $\mathcal{N}$ are the vertex set of a triangulation $\mathcal{T}$.

The integral of an extension of f over a triangulation gives a terrible approximation to $\int h\,d\chi$: an error of ± 1 on K nodes can cause a change in the integral of order K. More specifically, if $h = f + e$, where $e : \mathcal{N} \to \{-1, 0, 1\}$ is an error function that is nonzero on a sparse subset $\mathcal{N}' \subset \mathcal{N}$, then, for certain infelicitous choices of $\mathcal{N}'$, $|\int h - \int f| = |\mathcal{N}'|$.

A $\mathbb{R}$-valued relaxation can mitigate errors by using fluctuation c as a weight. Let $N(i)$ be a collection of neighboring nodes to x_i, where neighborhood can be defined via distance (if available) or edge-distance (in an ad hoc network or triangulation). Define $\tilde{h}$ to be the result of averaging the value at $x_i \in \mathcal{N}$ over $N(i)$, with c as a weight. Specifically,

$$\tilde{h}(x_i) = \frac{\sum_{y \in N(i)} c(y)h(y)}{\sum_{y \in N(i)} c(y)}. \tag{1.26}$$

This nearest-neighbor convolution damps out local variations. The resulting integral with respect to $\lfloor d\chi \rfloor$ will tend to mitigate localized errors, thanks to the Morse-theoretic formula.More numerical investigation is warranted.

Such averaging leads to non integer-valued integrands. By using integration with respect to $\lfloor d\chi \rfloor$ or $\lceil d\chi \rceil$ for upper/lower semicontinuous integrands associated to an averaged signal $\tilde{h}$, one obtains an expected value of $\int h\, d\chi$. This can be particularly illuminating when a network has incomplete information, *e.g.*, a hole. Holes in a network can be modeled by setting the confidence measure c to zero and averaging.

1.6.2. *Numerical integration.* Though integration with respect to Euler characteristic has a lengthy history, there appears to be no treatment of numerical integration, even in the simpler setting of $CF^+(\mathbb{R}^n)$. The central problem (in the constructible and definable categories) is how to estimate $\int_X h$ given the values of h on a discrete subset of X. As in the case of numerical integration for Riemann integrals, one typically assumes something about the features of h and/or the density and extent of the sampling set. In **[1]**, the present authors give a formula for estimating $\int h\, d\chi$ given a discrete sampling of $h \in CF(\mathbb{R}^2)$ which correctly samples connectivity data of excursion sets. This formula generalizes to the definable category:

PROPOSITION 23. *For $h \in \mathsf{Def}(\mathbb{R}^2)$ continuous, $\int h \lfloor d\chi \rfloor =$*

$$\int_{s=0}^{\infty} \beta_0\{h \geq s\} + \beta_0\{h \geq -s\} - \beta_0\{h < s\} - \beta_0\{h < -s\}\, ds, \tag{1.27}$$

where $\beta_0(\cdot) = \dim H_0(\cdot; \mathbb{R})$, the zeroth Betti number.

The value of Proposition 23 is that it allows for computation based on β_0 quantities. Such connectivity data are easily obtained from a discrete sampling via clustering. We have implemented this formula in software.However, for more general integration domains than $\mathbb{R}^2$, duality formulae are less helpful. One general result on refinement follows from continuity of the integral operator.

THEOREM 24. *For $h \in \mathsf{Def}(X)$ continuous, let h_{PL} be the piecewise-linear function obtained from sampling h on the vertex set of a triangulation $\mathcal{T}$ of X. As the sampling and triangulation are refined,*

$$\lim_{|\mathcal{T}| \to 0^+} \int_X h_{PL} \lfloor d\chi \rfloor = \int_X h \lfloor d\chi \rfloor. \tag{1.28}$$

This result relies crucially on continuity and does not apply to $CF(X)$. A more desirable result would be a measure of how far a given sampling is from the true integral. This seems challenging. We note that the Morse-theoretic formulae [1.16]-[1.17] allow one to reduce the domain of an integral to a (typically finite) set of critical points. This 'focusing' property of integration over Def should be a starting point for good numerical algorithms, especially for integral transforms.

2. Unimodal Category

2.1. Introduction. Here we introduce a novel topological approach to the problem of decomposing distributions $f : \mathbb{R}^n \to [0, \infty)$ into a convex combination of basis distributions. Instead of employing a decomposition into analytically defined (*e.g.*, normal) distributions, we propose a decomposition into topologically defined factors. Specifically, we consider the decomposition of a distribution into a sum of UNIMODAL distributions: those with a single maximum value and no other extrema. Such a decomposition is not uniquely defined; however, the minimal number of unimodal summands is. This UNIMODAL CATEGORY, $\mathtt{UCAT}(f)$, is a coarse measure of complexity for a distribution. The computation of UCAT is not obvious.

2.1.1. *Motivation.* There are several contexts within which the question about the number of summands in a unimodal decomposition arises. The first such context is, obviously, the statistical one. Indeed, unimodal distributions are the primal building blocks of statistical models. Essentially all classical probability distributions, including normal, Poisson, Gamma, Beta, Bernoulli, and more, are all unimodal. The methodology of statistical modeling essentially forces one to assume that the presence of several modes in a distribution is a consequence of its being a mixture of several distributions, and the relationship between the number of modes (essentially, the number of local maxima of the density, in the multivariate case) and the number of "components" of the mixture, *i.e.*, the number of summands in the convex decomposition, has been studied by many authors.

To justify the radical difference of our setup from the traditional statistical one — our lack of any assumptions about the structure of the summands beyond unimodality — we can invoke two considerations:

(1) The very variety of unimodal building blocks in standard statistical models suggests that one should try to abstract away any specific distribution, retaining only the minimal topological properties.
(2) Any specific analytic form of a density binds the distribution to some fixed coordinate system. Adopting a completely coordinate-free decomposition will inexorably relax to a topological approach.

We envisage applications of this topological decomposition to a variety of contexts within statistics, as well as data analysis and visualization (where decomposition can play a role in efficient encoding of the images or of multidimensional data).

The particular application inspiring this paper comes from signal analysis and reconstruction in a problem of target enumeration by means of a field of sensors [?]. In that setting, one has a finite number of targets located throughout a domain D. The domain is assumed to be filled with a network of primitive sensors. A sensor at $x \in D$ returns a number $h(x) \in \mathbb{N}$ which represents a count of how many targets it senses. The problem is to compute the number of targets given the function $h : D \to \mathbb{N}$. Assume each target α has a compact contractible *support* U_α on which it is observed by all sensors within. Then, the paper [?] computes the number of targets based on the Euler characteristics of excursion sets of h. When, however, one assumes a sensor modality which reads a $[0, \infty)$-valued signal instead of an integral count, the methods of [?] do not apply, and the techniques of this report take over. For example, in a network with acoustic or infrared sensors, it is more reasonable to expect sensors to return a real-valued signal which sums the contributions of in-range targets, these contributions varying in intensity as a function of range-to-target. Assuming a distribution of target sensory impact is a very weak assumption; however, the methods of this paper yield lower bounds on the enumeration problem.

2.1.2. *Statement of Results.* The extent of this report is to define the unimodal category, give a complete characterization and method of computation in the univariate case, and provide some key steps to understanding decompositions of higher-dimensional distributions. The specific contributions are as follow:

(1) We observe (Lemma 30) that the unimodal category is invariant under the right-action of a homeomorphism on the domain. This demonstrates the topological naturality of our definitions.
(2) We give a complete description (Theorem 36) of the unimodal category in the univariate case, along with a simple algorithm for its computation.
(3) We demonstrate that the univariate and bivariate cases, the unimodal number is a function of the Reeb graph of the distribution, labeled by critical values (Proposition 39). This implies that the unimodal category is purely a function of

Morse-theoretic data, providing a discretization of the problem.

(4) We prove a result (Proposition 41) about the 'persistence' of unimodal category with respect to translating the graph of the distribution. Shifting vertically can only decrease the unimodal category, and this number stabilizes to the value two.

(5) For distributions in arbitrary dimensions, we show (Proposition **??**) that a minimal unimodal decomposition exists with "depth" (the degree of overlaps of supports) bounded above by the dimension of the domain. We believe this result will be useful in computing the unimodal category by means of topological tools.

2.1.3. *Related work.* Morse structures associated to mixtures of multivariate normal distributions are discussed in **[32]**. This question already seems to be of interest to statisticians — even the simple univariate case is discussed in detail across several papers **[16, 4, 33]**, while the mixture of non-Gaussian unimodal densities is considered in **[24, 26]**. In particular, it is known that in mixtures of normal univariate distributions, the number of modes cannot exceed the number of components, a result which does not hold in higher dimensions (compare **[11]**).

The questions posed in this report are mostly complementary to the recent works **[9, 10, 12, ?, 30]** considering the estimation of the topology of data sets generated from noisy point-clouds. This is one set of works demonstrating the increasing interest in how topological and statistical phenomena entwine.

2.2. Unimodal category. The following definitions are common. We follow the notation of **[14]**.

DEFINITION 25. For X a topological space, the LUSTERNIK-SCHNIRELMANN CATEGORY of X, $\mathsf{LSCAT}(X)$, is the minimum number of open sets contractible in X which cover X. The GEOMETRIC CATEGORY of X, $\mathsf{gcat}(X)$, is the minimum number of open sets homotopic to a point which cover X.

Some authors (including those of **[14]**) use a reduced category, which measures the minimal number of open sets minus one. We do not follow this convention. Geometric category gcat is a homeomorphism invariant of a space, and L-S category LSCAT is a homotopy invariant. There are numerous deep connections between category and critical point theory (the classical motivation for the subject), dynamical systems, homotopy theory, and symplectic topology.

We introduce a new variant of category for distributions based on decomposition into unimodal factors.

DEFINITION 26. For X a topological space, let $\mathfrak{D} = \mathfrak{D}(X)$ denote the set of all compactly supported continuous distributions $f : X \to [0, \infty)$.

DEFINITION 27. A distribution $u \in \mathfrak{D}$ is said to be UNIMODAL if the upper excursion sets $u^c = u^{-1}([c, \infty))$ have the homotopy type of a point for all $0 < c \leq M$ and are empty for all $c > M$. Such a u has M as its maximal value.

We will refer to the nonempty upper excursion sets $u^c \subset X$ as being CONTRACTIBLE, though it must be clarified that such sets are contractible *in themselves* as opposed to contractible in X. The latter would be more in line with the definitions used in Lusternik-Schnirelmann theory, but would render the theory useless for most applications (where $X = \mathbb{R}^n$).

DEFINITION 28. Fix a norm $\nu = \|\cdot\|$ on $\mathbb{R}^N$. The UNIMODAL ν-CATEGORY of a distribution $f \in \mathfrak{D}(X)$ is defined as the minimal number $\mathtt{UCAT}^\nu$ of unimodal distributions $u_\alpha, \alpha = 1, \ldots, \mathtt{UCAT}^\nu$ on X such that f is pointwise the ν-norm of the collection (u_α). Specifically, $f(x) = \|(u_\alpha(x))\|$ for all $x \in X$.

The most natural and fundamental example is the minimal number of unimodal distributions required to represent a given distribution as a sum of unimodals. Summation of the components corresponds to the 1-norm on vectors, leading to the following generalization.

EXAMPLE 29. The UNIMODAL p-CATEGORY of a distribution $f \in \mathfrak{D}$ is the minimal number of unimodal distributions $u_\alpha, \alpha = 1, \ldots, \mathtt{UCAT}^p$ such that f is pointwise an ℓ^p combination. Specifically,

$$0 < p < \infty \quad : \quad f(x) = \left(\sum_\alpha (u_\alpha(x))^p \right)^{\frac{1}{p}} \tag{2.1}$$

$$p = \infty \quad : \quad f(x) = \max_\alpha \{u_\alpha(x)\} \tag{2.2}$$

LEMMA 30. *Any unimodal ν-category is invariant under the right-action of the homeomorphism group.*

2.3. Unimodal 1-category. Motivations, etc.

General observations.

REMARK 31. A trivial upper bound on $\mathrm{UCAT}^1(f)$ is the number of local maxima of f. The corresponding unimodal summands can be found supported on a small neighborhood of the basins of attraction of the gradient flow.

DEFINITION 32. Assume that a unimodal decomposition of $f = \sum_\alpha u_\alpha$ is given. A set $U \subset \sup(f)$ is called MAX-FREE if U does not contain any of the critical points of any f_α. For any max-free U we denote by $h(U)$ the DEPTH of U, the number of functions of the unimodal decomposition not vanishing on U:

$$h(U) = \#\{\alpha : U \cap \sup(u_\alpha) \neq \emptyset\}.$$

The following trivial lemma is useful:

LEMMA 33. *The depth of an open max-free set U is bounded from below by*

$$\left\lceil \frac{\max_U f}{\max_{\partial U} f} \right\rceil .$$

Solving the problem of computing unimodal 1-category is useful in other contexts:

LEMMA 34. *For all $f \in \mathfrak{D}$, $\mathrm{UCAT}^p(f) = \mathrm{UCAT}^1(f^{1/p})$.*

Univariate distributions. We commence with a computation of the unimodal 1-category on $\mathfrak{D}(\mathbb{R}^1)$. We assume that this function has isolated critical points. Up to a homeomorphism, this distribution is completely characterized by the *up-down* sequence of critical values corresponding to local minima and maxima,

$$0 = n_0 < m_0 > n_1 < \ldots < m_k > n_k = 0;$$

where $n_i = f(x_{2i})$; $m_i = f(x_{2i+1})$; $x_0 < x_1 < \ldots < x_{2k}$; and $i = 0, \ldots, k$, counting the initial and final point of the support as local minima.

PROPOSITION 35. *If an open interval (x_{2i}, x_{2j}) bounded by local minima is max-free (for some unimodal decomposition of f), then*

$$n_i - m_i + n_{i+1} - \ldots - m_{j-1} + n_j \geq 0. \tag{2.3}$$

Consider the open intervals with endpoints at the local minima $\{x_{2i}\}_1^k$. Call such an interval FORCED-MAX if the inequality (2.3) is violated there. Obviously, forced max intervals form an ideal: any interval containing a forced-max interval is itself forced-max.

THEOREM 36. *Let $f \in \mathfrak{D}(\mathbb{R})$ have maximal values $(m_i)_1^k$ and minimal values $(n_i)_0^k$ ordered according to the critical point order in the domain. Then $\mathtt{UCAT}^1(f)$ is equal to the maximal number of non-intersecting forced-max intervals:*

(2.4)

$$\mathtt{UCAT}^1(f) = \max\left\{N : (x_{2i_0}, x_{2i_1}), (x_{2i_1}, x_{2i_2}), \ldots, (x_{2i_{N-1}}, x_{2i_N})\,\textit{forced max}\right\}.$$

It is clear that for any collection of N nonintersecting forced-max intervals, the number of summands in the unimodal decomposition cannot be less than N: each forced-max interval, trivially, contains a critical point of at least one of the functions of the decomposition. The following algorithm yields an explicit unimodal decomposition *and*, simultaneously, a collection of nonintersecting forced-max intervals, one for each summand.

We construct the functions u_α iteratively, left to right, according to the algorithm Sweep. This entails sweeping f from the left and pulling out unimodal factors which, on their descent, compensate for the remaining factors as much as possible by descending according to the (positive) slope df. Here, the functions u_α are the differences of heights of the interleaving curves, which have the same differential as f on the intervals where f is nonincreasing. A new summand starts after the previous curve hits the x axis.

Algorithm 1 $\{u_\alpha\}$ = Sweep(f)

Require: $f \in \mathfrak{D}(\mathbb{R})$ with minima $n_i = f(x_{2i})_0^k$ and maxima $m_i = f(x_{2i-1})_1^k$

```
u_0 ⇐ 0 ; α ⇐ 1 ; g_α ⇐ f
while g_α ≠ 0 do
    y_α ⇐ first maximum of g_α from left
    u_α|(−∞, y_α] ⇐ g_α
    du_α|(y_α, ∞) ⇐ min(df, 0)
    u_α ⇐ max(u_α, 0)
    increment α
    g_α ⇐ f − Σ_{β<α} u_β
end while
return {u_β}_{β<α}
```

The proof that this construction is minimal follows from the observation that the local minima spanning the curves in the graphical construction form a max-forced partition of the support of f.

We have observed that, in the univariate case, the unimodal 1-category is a function of the critical values and the order in which they appear; however, there is not a strict dependence on this ordering. For example, in accordance with Lemma 30, UCAT^1 must be invariant under reversing the order of the critical points ($x \mapsto -x$). There is in fact a weaker dependence on the labeled Reeb graph of the distribution.

DEFINITION 37. The REEB GRAPH of a Morse function f on a smooth manifold M is the quotient space of M with respect to the relation declaring two points equivalent if they belong to the same connected component of a level set of f. In particular, the vertices of the graph are given by the connected singular level sets of f.

REMARK 38. In the univariate case, the Reeb graph is a rooted metric tree, *i.e.*, a rooted tree with lengths attached to its edges. The construction of this tree from an excursion is well-known in combinatorics and probability theory (cf. [31]). From our construction it follows that for a univariate distribution UCAT^1 depends only on its Reeb graph. In other words, one can swap the subexcursions bordering at a local maxima without affecting the unimodal number.

2.3.1. *Toward multivariate decompositions.* Multivariate distributions introduce a number of complexities. We sketch a few results in this section, reserving a more complete treatment for a later work.

An indication of the troubles arising in the multivariate case, we notice that an analogue of Proposition 35 is not valid there.

The dependence of the unimodal category on the Reeb graph noted in Remark 38 can be extended to planar distributions:

PROPOSITION 39. *The unimodal 1-category of $f \in \mathfrak{D}(\mathbb{R}^n)$ for $n = 1, 2$ is a function of the combinatorial type of the Reeb graph of f labeled by critical values.*

REMARK 40. In higher dimensions, the Reeb graph does not carry enough information: for example, a critical point of signature $(2, 1)$ is a vertex of valence 2 on the Reeb graph, with no information about the change of topology under the surgery defined by the point. In dimension 2, on the contrary, the change of topology is unambiguous.

Unimodal 1-category is not solely a function of the critical point and Morse index data of f, but rather depends upon the critical *values* as well. This dependence can be viewed in the following parameterized sense.

PROPOSITION 41. *Fix a (non-unimodal) C^1 distribution $f \in \mathfrak{D}(\mathbb{R}^n)$. For any $0 < p < \infty$, the unimodal p-category of the shifted distribution* *obtained by adding $C\mathbb{1}_{\sup f}$ to f for constant $C \geq 0$ is a non-increasing function of C which stabilizes to*

$$\lim_{C\to\infty} \mathtt{UCAT}^p(f + C\mathbb{1}_{\sup f}) = 1 + \mathtt{UCAT}^0(f) = 1 + \mathsf{gcat}(\sup f). \tag{2.5}$$

If one normalizes the distributions $f + C\mathbb{1}_{\sup f}$ to have unit mass, then one sees clearly the effect of increasing C is to reduce the total variation.

2.4. Unimodal ∞-category. Motivations, etc.

General observations. Computing the unimodal ∞-category in the univariate setting is trivial.

LEMMA 42. *For any $f \in \mathfrak{D}(\mathbb{R}^1)$, $\mathtt{UCAT}^\infty(f)$ equals the number of local maxima of f.*

LEMMA 43. $\mathtt{UCAT}^\infty$ *is invariant under the left action of* $\mathsf{Homeo}[0,\infty)$.

2.5. Bounds on unimodal category. In cases where exact computation of unimodal category is difficult, one turns to bounds based on analytic, geometric, or topological features of the distribution.

2.5.1. *Unimodal 0-category.* The monotonicity result allows one to define a unimodal category for ℓ^0.

DEFINITION 44. The unimodal 0-category, $\mathtt{UCAT}^0$, is defined as the limit of $\mathtt{UCAT}^p$ as $p \to 0$.

This limit is well-defined.

THEOREM 45. *For any $f \in \mathfrak{D}$, $\mathtt{UCAT}^0(f) = \mathsf{gcat}(\sup(f))$, the geometric Lusternik-Schnirelmann category of the support of f.*

PROOF. From Lemma 34, $\mathtt{UCAT}^p(f) = \mathtt{UCAT}^1(f^{1/p})$. As $p \to 0$, $f^{1/p}$ approaches $\mathbb{1}_{\sup(f)}$. $\square$

References

[1] Y. Baryshnikov and R. Ghrist, "Target enumeration via Euler characteristic integrals," *SIAM J Appl Math* 70(3), 2007, 825–844.

[2] Yu. Baryshnikov, R. Ghrist, "Euler integration over definable functions," *PNAS* 107, 2010, 9525–9530

[3] Yu. Baryshnikov, R. Ghrist, "Unimodal category", 2011, *submitted*.

[4] J. Behboodian, "On the modes of a mixture of two normal distributions," *Technometrics*, 12, 1970, 131–139.

[5] F. Bergeron, G. Labelle, and P. Leroux, *Combinatorial Species and Tree-like Structures*, Cambridge University Press, 1997.

[6] W Blaschke *Vorlesungen ber Integralgeometrie*, B.G. Teubner, Leipzig, 1936.

[7] G. Bredon, *Sheaf Theory*, 2^{nd} ed., GTM 25, Springer-Verlag, 1997.

[8] L. Brcker, "Euler integration and Euler multiplication," *Adv Geom* 5(1), 2005, 145–169.

[9] P. Bubenik and P. Kim, "A statistical approach to persistent homology," preprint (2006).

[10] G. Carlsson, T. Ishkhanov, V. de Silva, and A. Zomorodian, "On the local behavior of spaces of natural images," *Intl. J. Comput. Vision*, 76(1), (2008), 1–12.

[11] M. Carreira-Perpinan and C. Williams, "On the number of modes of a Gaussian mixture," in *Scale-Space Methods in Computer Vision, Lecture Notes in Comput. Sci.*, vol. 2695, 2003, 625–640.

[12] F. Chazal and A. Lieutier, "Weak feature size and persistent homology: computing homology of solids in $\mathbb{R}^n$ from noisy data samples," in *Proc. 21st Sympos. Comput. Geom.* (2005).

[13] R. Cluckers, M. Edmundo "Integration of positive constructible functions against Euler characteristic and dimension," *J Pure Appl Algebra* 208, 2008, 691–698.

[14] O. Cornea, G. Lupton, J. Oprea, and D. Tanré, *Lusternik-Schnirelmann Category*, Amer. Math. Soc., 2003.

[15] V. de Silva and R. Ghrist, "Coordinate-free coverage in sensor networks with controlled boundaries," *Intl. J. Robotics Research* 25(12), 2006, 1205–1222.

[16] I. Eisenberger, "Genesis of bimodal distributions," *Technometrics*, 6, 1964, 357–363.

[17] D. Estrin, D. Culler, K. Pister, and G. Sukhatme, "Connecting the Physical World with Pervasive Networks," *IEEE Pervasive Computing* 1:1, 2002, 59–69.

[18] K. Fukaya and Y.-G. Oh, "Zero loop open strings in the cotangent bundle and Morse homotopy," *Asian J. Math* 1, 99-180, 1997.

[19] K. Fukaya, Y.-G. Oh, H. Ohta and K. Ono, "Lagrangian intersection Floer theory-anomaly and obstruction," Kyoto University preprint, 2007.

[20] M. Goresky, R. MacPherson *Stratified Morse Theory*, (SpringerVerlag, Berlin), 1988

[21] H. Groemer, "Minkowski addition and mixed volumes," *Geometriae Dedicata* 6, 1977, 141–163.

[22] H. Hadwiger, "Integralstze im Konvexring," *Abh Math Sem Hamburg* 1956, 136154.

[23] T. He, P. Vicaire, T. Yan, L. Luo, L. Gu, G. Zhou, R. Stoleru, Q. Cao, J. Stankovic, T. Abdelzaher, "Achieving Real-Time Target Tracking Using Wireless Sensor Networks," in proceedings of *IEEE Real Time Technology and Applications Symposium*, 2006, 37–48.

[24] I. Kakiuchi, "Unimodality conditions of the distribution of a mixture of two distributions," *Math. Sem. Notes Kobe Univ.*, 9, 1981, 315–325.

[25] M. Kashiwara and P. Schapira, *Sheaves on Manifolds*, Springer-Verlag, 1994.

[26] J. Kemperman, "Mixtures with a limited number of modal intervals," *Ann. Statist.* 19, 1991, 2120–2144.

[27] D. Li, K. Wong, Y. Hu, and A. Sayeed, "Detection, classification, and tracking of targets," *IEEE Signal Processing Magazine*, 19(2), 2002, 17–30.

[28] R. MacPherson, "Chern classes for singular algebraic varieties," *Ann Math*, 1974, 423–432.

[29] J. Milnor, *Morse Theory*, Princeton University Press, 1963.

[30] p. Niyogi, S. Smale, and S. Weinberger, "Finding the homology of submanifolds with high confidence from random samples," *Disc. and Comput. Geom.*, to appear.
[31] J. Pitman, *Combinatorial Stochastic Processes.* Springer Lecture Notes in Mathematics, 1875, Springer-Verlag, Berlin, 2006.
[32] S. Ray and B. Lindsay, "The Topography Of Multivariate Normal Mixtures", *Annals of Statistics*, Vol. 33, No. 5, 2005, 2042–2065.
[33] C. Robertson and J. Fryer "Some descriptive properties of normal mixtures," *Skand. Aktuarietidskr*, 1969, 137–146.
[34] G.-C. Rota, "On the combinatorics of the Euler characteristic," *Studies in Pure Mathematics,* Academic, London, 1971, 221–233.
[35] P. Schapira, "Tomography of constructible functions," *11th International Symposium on Applied Algebra, Algebraic Algorithms, and Error-Correcting Codes,* 1995, 427–435
[36] P. Schapira, "Operations on constructible functions," *J Pure Appl Algebra* 72, 1991, 83–93.
[37] J. Schürmann, *Topology of Singular Spaces and Constructible Sheaves,* Birkhuser, Basel, 2003.
[38] O. Viro, "Some integral calculus based on Euler characteristic," Springer Lect Notes Math (Springer, Berlin), 1346, 1988, 127–138.
[39] L. Van den Dries *Tame Topology and o–Minimal Structures* Cambridge University Press, Cambridge, 1998.
[40] J. Yu and S. LaValle, "Tracking hidden agents through shadow information spaces," preprint, 2007.
[41] A. Zomorodian and G. Carlsson, "Computing Persistent Homology", *Disc. and Comp. Geom.*, 33, 2005, 249–274.